

Evidence gap map of performance measurement and management in primary healthcare systems in low-income and middle-income countries

Wolfgang Munar,¹ Birte Snilstveit,² Ligia Esther Aranda,³ Nilakshi Biswas,³ Theresa Baffour,³ Jenniffer Stevenson²

To cite: Munar W, Snilstveit B, Aranda LE, *et al*. Evidence gap map of performance measurement and management in primary healthcare systems in low-income and middle-income countries. *BMJ Global Health* 2019;**4**:e001451. doi:10.1136/bmjgh-2019-001451

Handling editor Seye Abimbola

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2019-001451>).

Received 30 January 2019
Revised 10 April 2019
Accepted 27 April 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Global Health, George Washington University Milken Institute of Public Health, Washington, District of Columbia, USA

²Synthesis and Reviews, 3ie London, London, UK

³Global Health, George Washington University Milken Institute of Public Health, Washington, District of Columbia, USA

Correspondence to
Dr Wolfgang Munar;
wolfgangmunar@gwu.edu

ABSTRACT

Introduction We mapped available evidence on performance measurement and management (PMM) strategies in primary healthcare (PHC) systems of low-income and middle-income countries (LMICs). Widely used, their effectiveness remains inconclusive. This evidence gap map characterises existing research and evidence gaps.

Methods Systematic mapping of performance measurement and management research in LMICs from 2000 to mid-2018; literature searches of seven academic databases and institutional repositories of impact evaluations and systematic reviews. Using a combination of manual screening and machine learning, four reviewers appraised 38 088 titles and abstracts, and extracted metadata from 137 impact evaluations and 18 systematic reviews that met the inclusion criteria. The resulting visual representation of the evidence base was uploaded to a web-based platform.

Results Since 2000, the number of studies has increased; the first systematic reviews were completed in 2010.

Two-thirds of the studies were conducted in sub-Saharan Africa and South Asia. Randomised controlled trials were the most frequently used study design. The evidence is concentrated in two types of PMM strategies: implementation strategies (in-service training, continuing education, supervision) and performance-based financing. Major gaps exist in accountability arrangements particularly the use of audit and feedback. The least studied types of outcomes were unintended effects, harm and social equity.

Conclusions The evidence is clustered around interventions that are unlikely to achieve transformational change in health outcomes. The gaps identified suggest that routinely used PMM strategies are implemented without sufficient knowledge of their effects. Future efforts at redesigning PHC systems need to be informed by evidence on the most effective approaches for using PMM strategies.

INTRODUCTION

Forty years after Alma Ata, the Astana Declaration of 2018 reaffirmed calls for placing primary healthcare (PHC) at the centre of

Key questions

What is already known about this topic?

- Performance measurement and management strategies that improve the performance of primary healthcare systems are central to the achievement of population-level health and equity outcomes in low-income and middle-income countries.

What are the new findings?

- Available evidence is relatively concentrated on provider-level interventions and strategies such as training and, to a lesser extent, supervision and performance-based financing; major gaps in evidence exist in accountability arrangements at organisational and community levels.

Recommendations for policy

- Efforts to strengthen or redesign primary healthcare systems need to be informed by new evidence on the most effective approaches for using performance measurement and management strategies.
- Future research needs to be theory driven and use mixed-methods designs to characterise what works or not, and why.

service delivery in the era of the Sustainable Development Goals (SDG) and universal health coverage.¹ The availability of high-performing PHC systems that serve as the first point of contact for the delivery of comprehensive, people-centred health services is an essential component for the achievement of these and other global health priorities. PHC systems also play an important role as a source of preparedness and response to disease epidemics and natural disasters.²⁻⁴ Furthermore, a recent report highlighted that quality-driven health systems could save up to 8 million lives annually in low-income and middle-income countries (LMICs) and that structural reforms are needed to improve the



Figure 1 Performance measurement and management (PMM) framework.

performance of individual providers, organisations and entire health systems.⁵

This paper reports one of several evidence gap maps (EGMs) commissioned to inform the launch of a PHC Research Consortium that delivers policy-relevant research in support of improving the performance and quality of health systems in LMICs.⁶ Informed by systematic approaches to evidence synthesis and review, EGMs were developed to systematically map evidence and research gaps on broad topic areas, describe the characteristics of the available evidence and inform the design of future research.⁷ In this paper, we summarise the framework and scope for a performance measurement and management (PMM) EGM, the methods and results, and implications for research.

PMM framework

For the purposes of this study, we defined PMM systems as the set of strategies, resources and capabilities for systematically measuring and improving the performance of healthcare delivery systems at the level of healthcare workers, patients, organisations and populations. The proximal aim of a functioning PMM system is to generate and sustain improvements in the behaviours and competences of healthcare workers and organisations alike, and increase the supply of high-quality services; their distal aim is to contribute to socially valued, population health and equity outcomes.

Based on this conceptualisation, we developed a PMM framework that integrates an existing public administration model of PMM systems⁸ and a framework developed by our team in the course of evaluating PMM strategies in LMICs.⁹ The resulting PMM system can be articulated as a series of iterative cycles with inter-related elements that include the policy and organisational context in which the healthcare system is embedded; cyclical measurement of performance at individual-wide, organisational-wide and system-wide levels; PMM strategies or interventions; the process of transformation of raw data into performance

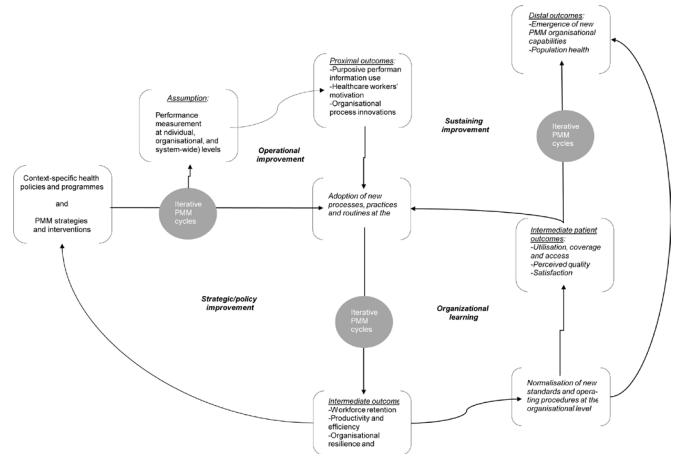


Figure 2 Performance measurement and management (PMM) theory of change.

information; sense-making and purposive use of performance information by health system actors; design and implementation of improvements; and the outcomes arising at the various levels of the PHC system (figure 1).

Drawing on organisational science, development economics, behavioural science, health systems research, public-sector administration and the sociology of organisations, we developed a high-level theory of how the PMM cycle described above may be operationalised to improve outcomes. Figure 2 maps the resulting theory of change of how PMM strategies can be linked to various outcomes.

The effective delivery and performance of a PHC system can be influenced by priority health policies and programmes, by PMM strategies and interventions, and by contextual conditions. The implementation of healthcare priorities and PMM strategies, and the use of iterative PMM cycles at multiple levels in a PHC system (individual, organisational and system-wide), can lead (or not) to proximal effects such as healthcare workers increased motivation to perform and to the use of performance information to experiment with operational innovations. The adoption (or not) of those innovations by system actors and the repetition of iterative cycles of measurement and process improvement (operational improvement loop in figure 2) can normalise the PMM cycles within the organisational culture and reinforce the adoption of new routines by growing numbers of system actors.¹⁰⁻¹²

The regular recurrence of improvement cycles through time may lead to the generation (or not) of intermediate outcomes like health workers’ retention, increased efficiency and productivity, among others, and to the emergence of new organisational capabilities for coordination of care, quality improvement, policy-making and regulation and others.¹³⁻¹⁶ Through time, these routines and practices can be further normalised and assimilated as new standards and operating procedures. The latter are necessary conditions for the generation of patient-level outcomes such as increased utilisation, effective

coverage, improved perception of quality and patient satisfaction. As patients' perceptions of quality increase, so can PHC service utilisation by households and communities. Assuming that the system's governance invests in the development of resources and capabilities to maintain gains in performance (organisational learning loop in figure 2), the recurrence of PMM cycles through time can lead (or not) to the questioning of long-standing practices and the emergence of new collective norms about patient and community care. The emergence of these capabilities for system learning and improvement are necessary conditions for the sustained delivery of high-quality PHC services and for PHC system performance (sustaining performance improvement loop in figure 2).¹⁶

As with any theory of change, a number of assumptions need to hold for PMM interventions to effectively change distal, population-level outcomes. For example, performance information needs to be available and used to trigger operational improvement or organisational learning. Also, the resulting performance management innovations need to be both well designed and implemented. In addition, PMM effects may be moderated by system antecedents and contextual factors, such as favourable organisational environments and leadership, and reforms to existing policies and regulations, among others.^{17–22} Ultimately, demand of quality services by households and communities is a necessary condition for PHC system performance. This is a function of socio-cultural and economic conditions, public policies such as use of incentives to households and/or healthcare providers, and the effective delivery of health promotion and outreach programmes, among others.

EGM scope

The above theory of change highlights the complexity of PMM strategies and the long and indirect causal chain from interventions to improved health outcomes. Mapping the available evidence and research gaps on PMM strategies in PHC systems also requires clear definitions of the interventions commonly used in practice and of its associated outcomes. To establish the scope of our EGM, we drew on an existing taxonomy developed by the Effective Practice and Organization of Care (EPOC).^{23–25}

We identified three categories of interventions: (1) implementation strategies; (2) accountability arrangements; (3) financial arrangements. Within these broad categories, we identified 15 PMM interventions that can operate at the individual, organisational or social levels. *Implementation strategies* are designed to bring about changes in the organisation of healthcare services, workers' behaviours or the use of health services by patients (eg, in-service training, supervision and continuous quality improvement, among others). *Accountability arrangements* are organisational, institutional and social arrangements used by health system actors for stewardship towards improved performance. Their focus can be internal or external. The former are focused on

reporting to and answering for the achievement of targets and milestones (eg, audit and feedback), while the latter relate to social systems in which citizens, consumers and/or communities exert control over the provision of public services (eg, social accountability and public release of performance information). *Financial arrangements* address performance by means of financial incentives that induce pro-performance behaviours among providers and healthcare organisations (eg, in-kind and financial incentives and pay-for-performance).

PMM interventions aim to improve outcomes at individual, organisational and collective levels. Changes may occur across short and long timeframes, and can include both desirable and undesirable adverse effects. We also defined five broad outcome categories, with several specific outcomes associated with each category. Table 1 describes the intervention and outcome categories included in this paper.

METHODS

This study systematically characterised the research done of the interventions and effects of PMM strategies and identified gaps in evidence. This EGM covers a broad body of literature covering PMM strategies that aim to change supply-side and, to a lesser extent, demand-side approaches like social accountability. Our methods differ from systematic reviews that usually address more narrow research questions. In completing this research, we followed the evidence-mapping methods developed by the International Initiative for Impact Evaluation (3ie).^{26–27} Such methods were outlined in detail in the published study protocol⁷ and are summarised below.

We systematically mapped the evidence on PMM strategies in PHC systems in LMIC settings that were published in any language between 1 January 2000 and 30 June 2018. Based on the theory of change and the taxonomy described above, we developed inclusion and exclusion criteria, and designed a comprehensive search strategy. The search strategy was developed to identify studies in Medline (PubMed), Embase (Ovid), CAB Global Health (Ovid), CINAHL (Ebsco), Cochrane Library, Scopus (Elsevier) and Econlit (Ovid). We also searched institutional databases and repositories of impact evaluations and systematic reviews. Search strategies are presented in online supplementary file 1.

We included studies in LMICs, as defined by the World Bank, designed to assess the interventions and outcomes listed in table 1 and systematic reviews of such studies. We used the PROGRESS-Plus framework to assess the extent to which intervention outcomes addressed equity and social differentiation into categories such as place of residence, race/ethnicity, occupation, gender, religion, education, socioeconomic status and social capital, age and disability.²⁸ We developed coding tools to extract meta-data from all impact evaluations included and appraised the methodology of systematic reviews using a standardised checklist²⁹ (online supplementary file 2).

Table 1 PMM intervention and outcomes included in the EGM

Interventions		Outcomes	
Implementation strategies at individual provider level	In-service training; continuous education; reminders	Provider and managerial outputs and outcomes	Workload; work morale; stress, burnout and sick leave; staff turnover; provider knowledge; change in attitudes, beliefs, perceptions; skills and competencies
Implementation strategies at organisational level	Supervision; clinical incident reporting; local opinion leaders; continuous quality improvement; clinical practice guidelines	Organisational outcomes	Quality of care process improvement; adherence to recommended practice or guidelines; patient satisfaction; perceived quality of care; changes in organisational culture
Accountability arrangements	Audit and feedback (at provider or organisational levels); public release of performance information; social accountability	Patient outcomes	Health behaviours (adherence by patients to treatment or care plans, and health-seeking behaviours); health status outcomes (physical health, or psychological and psychosocial outcomes)
Financial arrangements	Performance-based financing (at provider or organisational levels)	Health outcomes at population level	Utilisation of services; service or intervention coverage; access to primary care services; adverse health effects or harm
		Social and equity outcomes	Community participation; equity effects; unintended outcomes

EGM, evidence gap map; PMM, performance measurement and management.

The latter assesses the methods used to identify, include, appraise and synthesise studies in the review, and gives an overall rating to each review of high, medium or low confidence in review methods.

Included impact evaluations used experimental or observational data to measure the effects of a programme relative to a counterfactual. We included randomised controlled trials (RCTs), regression discontinuity designs, controlled before-and-after studies using appropriate methods to control for selection bias and confounding, instrumental variables estimation or other methods using an instrumental variable, and difference-in-difference study designs. Cross-sectional or panel studies were included that had an intervention and comparison group, and interrupted time series studies that used at least three observations before and three after the intervention. Studies in high-income countries, efficacy trials and systematic reviews of efficacy trials were excluded. We used an online systematic review application (EPPI-Reviewer V.4) to manage the literature database, screen for eligibility and extract data³⁰ and developed and refined data extraction coding tools.

Four reviewers were trained on screening, eligibility criteria and data extraction. All reviewers first screened the same set of 100 studies, and we discussed and resolved any disagreements between reviewers. This process ensured a common interpretation of inclusion criteria within the team. In addition, we also operated with a safety-first approach³¹ whereby any studies where the first screener was uncertain about inclusion/exclusion were screened by a second, more senior reviewer.

A total of 38 088 articles were uploaded unto EPPI-Reviewer including 38 056 from the systematic searches and 32 from the grey literature and citation tracking. From this original set, we manually excluded 6277 records (16.5%); an additional 21 543 (83.5%) were excluded using EPPI-Reviewer's machine-learning capabilities, a text-mining feature that allows automatic term recognition and document clustering.³² The resulting 555 studies were screened at full text, after which 200 were excluded with reasons. The most common reasons for exclusion were study design and intervention criteria. The final EGM included 137 impact evaluations and 18 systematic reviews. [Figure 3](#) summarises the screening process.

We conducted descriptive analysis of the database of literature included in EPPI-Reviewer V.4 across geography, time, interventions and outcomes, study design and equity considerations. Resulting data were uploaded to 3ie's website to create a visual representation of the evidence.

Patient and public involvement statement

Patients were not involved in the design or implementation of this study nor will results be disseminated among patients. However, the research question and the resulting EGM highlight patient-level interventions and outcomes.

RESULTS

Between 2000 and mid-2018, the total number of studies in the PMM field increased, from 2 in 2000 to 23 in

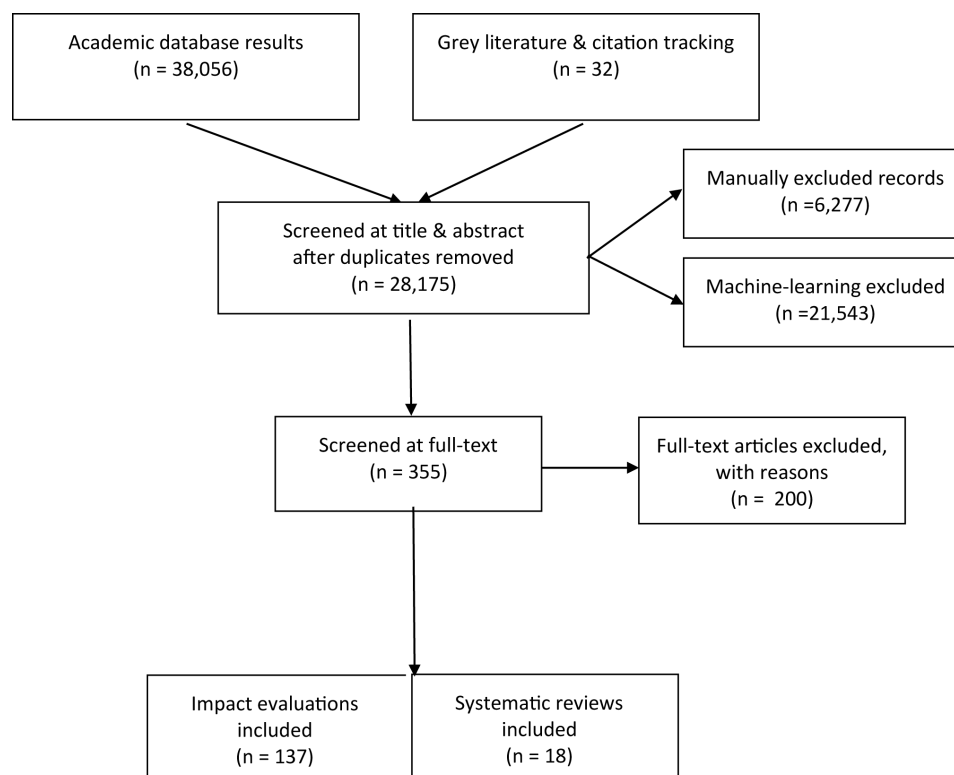


Figure 3 Screening process.

2017; impact evaluations were conducted in 42 LMICs; two-thirds of the studies were completed in Sub-Saharan Africa (SSA) (56%) and South Asia (19%); no systematic reviews were published prior to 2010. RCTs were the most frequently used study design (71%), followed by controlled before-and-after studies (21%). Thirty studies adopted mixed-methods study designs. The most common methods of analysis were multivariate regression, difference-in-difference estimation, and comparison of means. Few studies used fixed-effects or random-effects models, instrumental variables or propensity score matching. RCTs were less frequent in Latin America and the Caribbean (LAC).

The most frequently studied PMM approaches were implementation strategies such as in-service training, supervision and continuous education, and on financial arrangements such as performance-based financing (PBF). These types of strategies accounted for more than 90% of studies in SSA and LAC, more than two-thirds in South Asia and Middle East and North Africa (MENA), and 65% in East Asia and the Pacific (EAP). PBF and incentives were under-represented in South Asia (4%) and EAP (12%). We identified a small cluster of continuous quality improvement studies in SSA (5%) while all other regions showed absolute gaps in this type of intervention. Similarly, supervision interventions were studied in South Asia (31%), SSA (8%) and EAP (6%), but no studies were found in LAC or MENA. While the category of accountability arrangements was the least studied overall, 24% of studies in South Asia addressed public release of performance information, and 8% in

SSA evaluated social accountability interventions. Other than a few studies in SSA (2%), the largest gaps were in the area of audit and feedback.

The most frequently assessed outcomes were effects at the provider level, followed by patient-level and population-level health. Among the former, the acquisition of knowledge and skills and adherence to guidelines were the most frequently measured effects; the least assessed were work morale; providers' attitudes, beliefs and perceptions; turnover and retention; workload; stress; burnout; and sick leave. Patient behavioural and mental health outcomes and adherence to treatment were infrequently studied. Two-thirds of studies addressed more than one outcome.

At the organisational level, the most frequently addressed outcomes were adherence to recommended practice or guidelines, process of care improvements, patient satisfaction and perceived quality of care. Population-level outcomes included utilisation or coverage of health services, and, less frequently, access to services. Overall, the least studied outcomes were unintended effects, harm, and social and equity outcomes. No studies measured changes in organisational culture.

Twenty-eight impact evaluations addressed equity to some extent. The majority of studies analysed subgroups or targeted populations. Very few studies considered an equity-sensitive research question, assessed equity outcomes, adopted an equity-sensitive method to assess how and why effects were produced, or followed an equity-informed research process. No study adopted equity-sensitive analytical frameworks or theories of

change. Among the studies that addressed equity, the most commonly addressed dimensions were place of residence and socioeconomic status. Few studies considered other dimensions such as gender; vulnerable groups; race, ethnicity, culture and language; education; or age. No studies considered disability, occupation, religion or social capital. Equity considerations included in the systematic reviews were almost absent.

Although we only found 18 systematic reviews, their quality was reasonably high. Nine reviews were rated as high confidence, two as medium confidence and the remaining seven as low confidence. Of the high-confidence reviews, five were published in the Cochrane Collaboration Library, three were published in peer-reviewed journals and one in the Campbell Collaboration Library. Reviews rated as high confidence had clear inclusion criteria, conducted comprehensive searches that included sources of grey and unpublished literature, undertook independent double-screening of included studies, and undertook and clearly reported risk of bias assessments using sensible criteria. They also typically undertook at least partial independent double data extraction, clearly reported the characteristics and results of included studies, combined the results of included studies appropriately and explored heterogeneity in results, including by risk of bias of the included studies where appropriate. The reasons for the systematic reviews being downgraded to low confidence included a combination of conducting a limited search that did not include grey literature, selective reporting of results of included studies, lack of independent screening of full-text papers, not assessing the risk of bias or quality of included studies, and using vote counting by direction of effect and statistical significance as a method of synthesising findings.

Reflecting the IE evidence base, systematic reviews also focused on individual-level PMM interventions such as in-service training, continuous education and supervision. Nine systematic reviews addressed PBF, six of which focused on healthcare workers' behaviours and three on organisational-level change. Several reviews considered more than a single intervention. Specific outcomes addressed included adherence to guidelines and provider knowledge and skills. Seven systematic reviews addressed patient-level mortality. Population-level outcomes included increased access, coverage and utilisation of specific services along with changes in patient knowledge, adherence, satisfaction and perceived quality of service. Only one review measured unintended patient, organisational, population and social outcomes, and two more addressed changes in work morale and provider attitudes, beliefs and perceptions. Three of the 18 systematic reviews had an equity focus.

A summary of results is presented in [table 2](#). The resulting map (not included here due to its size) is available in 3ie's website (<http://gapmaps.3ieimpact.org/evidence-maps/performance-measurement-and-management-primary-care-delivery-systems>). In the

map, readers will find interventions on the y-axis and outcomes on the x-axis. Included studies are presented in the form of bubbles according to the interventions and outcomes addressed. Bubbles indicate the relative size of the evidence base; grey bubbles show impact evaluations and coloured bubbles identify systematic reviews and protocols. The confidence rating assigned to the systematic reviews in the critical appraisal is also indicated by colours (red, low confidence; orange, medium confidence; green, high confidence). Linkages to all studies are available online and are listed in online supplementary file 3.

DISCUSSION

We conducted an EGM of the evidence on the effects of PMM system interventions in PHC systems in LMICs to characterise existing research and identify evidence gaps. Overall, our main finding is that at a global level, the evidence is concentrated on implementation strategies such as in-service training and continuing education, and financial arrangements such as PBF and financial incentives. More specifically, there are PMM strategies for which we found few or no studies. These absolute gaps are concentrated in the area of accountability arrangements and, specifically, on interventions such as audit and feedback, public release of performance information and social accountability. In the category of implementation strategies, there was an absolute gap in the area of continuous quality improvement.

To date, the evidence base is mainly focused on individual-level outcomes among healthcare workers such as the acquisition of knowledge and skills and the adherence to practice guidelines. Patient-level outcomes are also frequently measured, in particular the utilisation and coverage of specific services or biomedical interventions. Globally, gaps in outcomes are concentrated on organisational-level and on the social-level and equity-level effects of most PMM strategies and interventions. There is also an almost-complete absence of studies that address unintended and harmful effects. The patterns identified for impact evaluations are largely mirrored in the systematic reviews available.

Research to date has prioritised the evaluation of PMM strategies that appear to operate, in the short term, through changing the behaviours of healthcare workers and patients. The underlying programme theories appear to assume that changing PMM inputs and financing processes would somehow ensure population-level health outcomes. However, our theoretical framework suggests that PMM strategies behave like complex systems whose functioning requires coordinated and sustained efforts operating at multiple levels going beyond improving individual providers' skills, knowledge or incentives. Implementation of under-theorised PMM strategies and interventions may contribute to this, but the prevalence of 'single method' studies originating from a single discipline is also likely to be an important factor. We found

Table 2 PMM strategies addressed in the literature between 2000 and 2018: regional distribution, study designs, intervention strategies and outcomes

	Sub-Saharan Africa (col. %)	South Asia (col. %)	East Asia and Pacific (col. %)	Latin America and Caribbean (col. %)	Middle East and North Africa (col. %)
Impact evaluations	56	19	11	9	5
Study designs					
RCT	73	81	75	54	63
Controlled before and after	19	15	19	38	13
Other designs per protocol	8	4	6	8	24
PMM strategies and interventions					
1. Implementation strategies					
In-service and continuing education	41	50	47	69	63
Supervision	8	31	6	–	–
Continuous quality improvement	5	–	–	–	–
2. Accountability arrangements					
Social accountability	8	–	–	–	–
Public release of performance information	–	24	–	–	–
Audit and feedback	2	–	–	8	–
3. Financial arrangements					
Performance-based financing and incentives	27	4	12	23	25
Outcomes					
Provider outputs and outcomes	32	54	61	28	56
Organisational outcomes	17	3	11	11%	11%
Patient outcomes	14	43	–	22%	33%
Health outcomes at population level	33	–	28	39%	–
Social and equity outcomes	4	–	–	–	–

PMM, performance measurement and management; RCT, randomised controlled trial.

that a majority of the identified impact evaluations adopted experimental designs, and while such studies generate reliable evidence on effects, the infrequent use of mixed-method impact evaluations also means that important research questions such as how and why change happens (or not), for whom and at what costs are not addressed.^{33 34}

Given the strategic role that PHC systems can play in ensuring that differentials in health outcomes are reduced, the lack of attention to equity in the evidence base is a major concern as countries face the challenge of attaining the SDGs. Furthermore, unsafe practices, harmful effects and unintended outcomes have not been systematically addressed in the literature. This is also relevant for the SDG agenda, given that PMM interventions can create perverse incentives and undesirable effects with potential for worsening or further deepening existing inequities.

Despite differences in aim and scope, a recent systematic review (not included in our dataset) of the effectiveness of strategies to improve the practice of healthcare

providers in LMICs³⁵ reached similar conclusions. In particular, the authors found that the strategies used had low-quality evidence and that comparisons between studies were difficult given contextual and methodological heterogeneity. Our findings also echo those reported by Shojania and Grimshaw more than a decade ago in their review of the evidence on quality improvement interventions.³⁶ Similar to their findings, we find a lack of convincing theories underlying the selection and design of PMM strategies studied in the literature. Furthermore, much of the literature fails to meet commonly accepted standards for the conduct and reporting of research. Evidence on effects is necessary, but not sufficient, for informing policy and practice. The field requires approaches to impact evaluation that are informed by well-articulated theory and that effectively integrate methods for assessing effects with qualitative and systems science approaches that are better suited to address the interactions between context and implementation processes.^{37 38}

Implications for research

Addressing the evidence gaps identified in this study requires researchers, commissioners of research and decision-makers to get behind an ambitious research agenda to ensure efforts to strengthen PHC systems that are informed by evidence on the most effective approaches for using PMM strategies. To ensure the relevance, uptake and use of such new evidence to inform policy and practice, a strategic research agenda should be informed by the findings of this and other EGMs and co-developed with researchers, policy-makers and practitioners in LMICs.

The underlying causes for the similarities and differences in the evidence base across geographical regions need to be further analysed. Consideration should be given to the role of funders, official development assistance agencies and researchers in shaping the current evidence base. The preference for experimental designs may signal that research capacity has expanded in important aspects. However, evidence of duplicative efforts in the area of in-service training, the focus on 'micro' or individual-level outcomes, and the absolute gaps identified in certain types of interventions and outcomes may reflect lack of capacity for conducting organisational and multi-level systems and policy research, a need that should be addressed by future research efforts.

In particular, we believe that there is a need for research that tests the theory of change and the associated assumptions underlying the implementation of complex, adaptive PMM strategies. Instead of focusing on assessing the effectiveness of specific, discrete health interventions as has been the trend to date, PMM systems theory suggests transformational change can only be achieved by multilevel interventions designed to trigger changes in provider, patient and organisational-level behaviours. The specific form of such PMM interventions are likely to vary across contexts and might be mediated by conditions such as domestic implementation capabilities, policy and organisational environments, and other 'binding constraints'.³⁹ Therefore, a necessary condition for developing a rigorous evidence base is ensuring that future research assesses the effects of programmes that have been developed based on an appropriate contextual diagnosis and that include interventions that target the key constraints identified in such analyses.

From these broader considerations, the following implications emerge:

- ▶ There is a need for a coordinated research and learning agenda organised around a common multidisciplinary conceptual framework of PMM strategies as complex adaptive systems.
- ▶ The selection of specific interventions to be tested should be informed by end-user needs and diagnosis of binding constraints in specific health systems.
- ▶ Studies should integrate a strong theory of change with rigorous empirical investigations.
- ▶ Studies should respond to the range of questions that need to be answered to effectively inform the

development and implementation of new policies and programmes, including what works, why (or why not), for whom and at what cost, including by (1) effectively integrating quantitative and qualitative research methods, for example, by combining study designs that allow for causal inference with process evaluations, and qualitative research; (2) considering both intended and unintended effects; and (3) considering effects on disadvantaged or vulnerable groups through equity-sensitive study designs.^{40 41}

- ▶ To improve value for money of new studies, researchers and commissioners should ensure that future research adopts commonly accepted standards for research transparency, including pre-registration, sharing of data on study completion and comprehensive study reporting.⁴²

Study strengths and limitations

This study has several strengths. We used a multidisciplinary conceptual framework for defining PMM, drawing from several disciplines to systematically search seven databases, thus providing a richer picture of PMM PHC systems in LMIC. Such theory of change allowed us to study PMM interventions and outcomes using a theorised, multilevel perspective of individual, organisational and system-wide levers for transformational change. Finally, we followed a methodological approach for the design of EGMs, including a systematic consideration of equity.

Due to resource constraints, several limitations must be acknowledged. First, the scope for this EGM mainly addressed the supply side of publicly delivered PHC services, therefore providing a partial view of the landscape, particularly in view of comprehensive perspectives of PHC systems that include private provision, demand-side perspectives and rights-based approaches.

We used the EPOC taxonomy for defining interventions and outcomes as these categories are well understood in the health systems and health services research community. However, this meant, first, that evidence originating in disciplines other than health services and health system research may not be part of the EPOC taxonomy and, therefore, will have been excluded from the EGM. Additionally, operational constraints did not allow for double screening of studies at the title and abstract stage. However, we implemented several strategies to ensure the quality of the screening process as described above.

Finally, hospital-based interventions were excluded as non-PHC interventions. However, primary care maternal and child health services, particularly deliveries, are provided in hospital settings in some countries. Therefore, by excluding hospital care, we may have excluded some studies of services that are technically considered part of PHC.

CONCLUSIONS

In view of the recent Astana Declaration¹ and calls for high-quality health systems for achieving the SDGs,⁵ the limited evidence available on PMM strategies is a major concern. The evidence that does exist is clustered around relatively 'simple' interventions that are unlikely to achieve the transformational change in health outcomes that will be required to reach the SDGs and save and improve the lives of the millions of people without access to high-quality health services, particularly in LMICs. The evidence gaps identified suggest that routinely used PMM strategies are being implemented without sufficient knowledge of their effects, intended and otherwise. We call on researchers, commissioners and decision-makers to get behind an ambitious research agenda which can ensure future efforts at strengthening or redesigning PHC systems are informed by evidence on the most effective approaches for using PMM strategies to improve outcomes.

Acknowledgements Preliminary results of this study were presented at a satellite session of the 5th Global Health Systems Research Symposium, on 6 October 2018. We acknowledge support for this study from the Salud Mesoamerica Initiative and the Inter-American Development Bank. We also thank Emma Iriarte, Pedro Bernal Lara, Jennifer Nelson, Ignez Tristao, Maria D Sanchez and Elizabeth Bastias-Butler for their inputs.

Funding This work was supported by the Bill and Melinda Gates Foundation through grants to George Washington University (OPP1154415) and Ariadne Labs through Brigham and Women's Hospital (OPP1149078).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All data relevant to the study are included in the article or uploaded as online supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. WHO. Declaration on primary health care. Astana, 2018, 2018. Available: <http://www.who.int/primary-health/conference-phc/declaration>
2. Bitton A, Ratcliffe HL, Veillard JH, *et al*. Primary health care as a foundation for strengthening health systems in low- and middle-income countries. *J Gen Intern Med* 2017;32:566–71.
3. Gates B. The next epidemic—lessons from Ebola. *N Engl J Med* 2015;372:1381–4.
4. Lewin S, Lavis JN, Oxman AD, *et al*. Supporting the delivery of cost-effective interventions in primary health-care systems in low-income and middle-income countries: an overview of systematic reviews. *Lancet* 2008;372:928–39.
5. Kruk ME, Gage AD, Arsenault C, *et al*. High-quality health systems in the sustainable development goals era: time for a revolution. *Lancet Glob Health* 2018;6:e1196–252.
6. Hirschhorn L, Langlois E, Ghaffar A. Alma Alta at 40: what do we still need to know. *BMJ Global Health* 2019.
7. Munar W, Snilstveit B, Stevenson J, *et al*. Evidence gap map of performance measurement and management in primary care delivery systems in low- and middle-income countries – study protocol. *Gates Open Res* 2018;2.
8. Pollitt C. Performance management 40 years on: a review. Some key decisions and consequences. *Public Money & Management* 2018;38:167–74.
9. Munar W, Wahid SS, Curry L. Characterizing performance improvement in primary care systems in Mesoamerica: a realist evaluation protocol. *Gates Open Res* 2018;2.
10. Berwick DM. The science of improvement. *JAMA* 2008;299:1182–4.
11. Lemire S, Christie CA, Inkelas M. The methods and tools of improvement science. *New Dir Eval* 2017;2017:23–33.
12. May CR, Mair F, Finch T, *et al*. Development of a theory of implementation and integration: normalization process theory. *Implement Sci* 2009;4.
13. Eden C, Ackermann F. Mapping distinctive competencies: a systemic approach. *J Oper Res Soc* 2000;51:12–20.
14. Eden C, Competences AF. Distinctive competencies, and core competencies. *Research in Competence-Based Management* 2010;5:3–33.
15. Piening EP. Dynamic capabilities in public organizations: a literature review and research agenda. *Public Management Review* 2013;15:209–45.
16. Nembhard IM. Learning and improving in quality improvement collaboratives: which collaborative features do participants value most? *Health Serv Res* 2009;44:359–78.
17. Wilkens U, Menzel D, Pawlowsky P. Inside the black-box: analysing the generation of core competencies and dynamic capabilities by exploring collective minds. An organisational learning perspective. *management revu* 2004;15:8–26.
18. Kok MC, Kane SS, Tulloch O, *et al*. How does context influence performance of community health workers in low- and middle-income countries? Evidence from the literature. *Health Res Policy Syst* 2015;13.
19. Theobald S, Hawkins K, Kok M, *et al*. Close-to-community providers of health care: increasing evidence of how to bridge community and health systems. *Hum Resour Health* 2016;14.
20. Kok MC, Broerse JEW, Theobald S, *et al*. Performance of community health workers: situating their intermediary position within complex adaptive health systems. *Hum Resour Health* 2017;15.
21. Greenhalgh T, Robert G, MacFarlane F, *et al*. *Diffusion of innovations in health service organisations: a systematic literature review*. Malden. 581. MA: Blackwell Publishing, 2005.
22. Piening EP, Salge TO. Understanding the antecedents, contingencies, and performance implications of process innovation: a dynamic capabilities perspective. *J Prod Innov Manage* 2015;32:80–97.
23. Care) ECEPaOo. *EPOC taxonomy*. The Cochrane Collaboration, 2015.
24. Care) ECEPaOo. Describing interventions in EPOC reviews 2017, 2018. Available: <http://epoc.cochrane.org/resources/epoc-resources-review-authors> [Accessed Feb 1, 2018].
25. Care) ECEPaOo. What outcomes should be reported in EPOC reviews?: Cochrane Collaboration 2017. Available: <http://epoc.cochrane.org/resources/epoc-resources-review-authors> [Accessed 1 Feb 2018].
26. Snilstveit B, Vojtkova M, Bhavsar A, *et al*. *Evidence gap maps—a tool for promoting evidence-informed policy and prioritizing future research*. Policy research Working papers. Washington, DC: The World Bank, 2013.
27. Snilstveit B, Bhatia R, Rankin K. *3ie evidence gap maps. a starting point for strategic evidence production and use*. 3ie Working papers. 24. New Delhi, India: International Initiative for Impact Evaluation, 2017.
28. O'Neill J, Tabish H, Welch V, *et al*. Applying an equity lens to interventions: using progress ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014;67:56–64.
29. Snilstveit B, Eyers J, Bhavsar A, *et al*. *3ie database of systematic reviews in international development: search strategy and procedures document*. London, UK: International Initiative for Impact Evaluation, 2014.
30. Thomas J, Brunton J. *EPPI-Reviewer: software for research synthesis*, 2007.
31. Shemilt I, Khan N, Park S, *et al*. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. *Syst Rev* 2016;5.
32. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4: software for research synthesis*. London: UK: EPPI-Centre Software Social Science Research Unit, UCL Institute of Education, 2018.
33. Pawson R. *Evidence-based policy: a realist perspective*. Thousand Oaks, CA: Sage Publications, 2006.
34. Pawson R, Tilley N. *Realistic evaluation*. Sage, 1997.

35. Rowe AK, Rowe SY, Peters DH, *et al.* Effectiveness of strategies to improve health-care provider practices in low-income and middle-income countries: a systematic review. *Lancet Glob Health* 2018;6:e1163–75.
36. Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. *Health Aff* 2005;24:138–50.
37. Williams B. Prosaic or profound? The adoption of systems ideas by impact evaluation. *IDS Bull* 2015;46:7–16.
38. Williams B, Imam I. *Systems concepts in evaluation: an expert anthology*. Point Reyes, CA: EdgePress of Inverness, 2006.
39. Gaarder M, Dixon V. Misdiagnosis and the evidence trap: a tale of inadequate program design: 3Ie, 2018/2018. Available: <http://www.3ieimpact.org/blogs/misdiagnosis-and-evidence-trap-tale-inadequate-program-design> [Accessed 18 Dec 2018].
40. Leach B, Snilstveit B. *3ie gender and equity coding tool*. Delhi/ London: International Initiative for Impact Evaluation, 2018.
41. Welch VA, Norheim OF, Jull J, *et al.* CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. *BMJ* 2017;359.
42. 3ie Transparency policy Delhi/London. International initiative for impact evaluation (3Ie), 2018. Available: <http://www.3ieimpact.org/sites/default/files/2018-05/3ie-policy-research-transparency-policy%20vApr18.pdf> [Accessed Jan 4, 2019].